DOCUMENT RESUME

ED 257 669 SE 045 765

AUTHOR Gabb, Roger G.
TITLE A Comparison of Two Uses of Video in the Evaluation of Laboratory Programs.
PUB DATE Dec 84
NOTE 23p.; Paper presented at the Annual Meeting of the American Vocational Association Convention (New Orleans, LA, November-December, 1984).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; *College Science; Comparative Analysis; Foreign Countries; Higher Education; *Interaction; Program Evaluation; Science Activities; Science Education; Science Instruction; *Student Behavior; *Student Reaction; *Videotape Recordings
IDENTIFIERS *New Zealand; Science Education Research

ABSTRACT

Two different methods of collecting information about student activity in the laboratory were compared using a series of videotape recordings. One method was based entirely on the systematic observation of student behavior (interaction analysis); the second method involved students describing their laboratory activities while watching excerpts from the videotape recording of the laboratory session (stimulated recall). Subjects, first-year undergraduate students taking 4-year courses in agricultural and horticultural science at Lincoln College (New Zealand), were selected by teachers to participate in the study. The activity of a single student in each of 13 laboratory classes (5 in zoology and 8 in physics) was recorded to compare the two methods. To determine the relative utility of the information produced by the two methods (including usefulness of the information, preference for one or the other method, and limitations of information provided), nine staff members were given a brief introduction to the project and summaries for the two methods. All nine teachers were interviewed and indicated that the information was useful in developing laboratory programs. In addition, they expressed a preference for the information generated by the second method (stimulated recall). (JN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED257669

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

A Comparison of Two Uses of Video in the Evaluation of Laboratory Programs.

Roger G. Gabb, B.V.Sc., Ph.D. (Sydney), M.Ed. (Bath)
Director
Educational Services Unit
Lincoln College
Canterbury
NEW ZEALAND.

Paper presented at the 1984 American Vocational Association Convention (AVERA Research Section), New Orleans, November - December 1984.

SE 045 765

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Roger G. Gabb
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Hundreds of hours are devoted to laboratory programs in most college-level applied science courses and there is therefore a substantial commitment of both human and material resources in their support. However, despite their costly nature and their almost universal use in science and technology courses, there has been relatively little scrutiny of the educational effectiveness of laboratory programs until recently. The recent growth of interest in the evaluation of laboratory work probably reflects, at least in part, the growing concern of many college administrators with the cost-effectiveness of this relatively expensive method of teaching.

One approach to the improvement of the effectiveness of laboratory programs is to study the relationship between student activity in class (the process of laboratory work) and what students learn from the class (the learning outcomes of laboratory work). However, this approach presents some difficulties. Not only is it difficult to measure some learning outcomes of laboratory work but there is also little agreement on effective methods of collecting information about student activity in class (Power, 1977).

While many writers on evaluation advocate the collection of information about student activity in class (eg. Brinkerhoff, Brethower, Hluchyj and Nowakowski, 1983), they do not agree on what approach should be used to do so. Some favour approaches based on interaction analysis or the coding of student behavior in the class (Shymansky and Penick, 1979) while others prefer to rely on student self-reports of their activity or unstructured observation (Ogborn, 1977).

In this paper, we report on a methodological study in which we compared two quite different methods of collecting information about student activity in the laboratory. One method was based entirely on the systematic observation of student behavior while the other focused on student description of their activity in the laboratory.

PROCEDURE

To make a valid comparison between two methods of studying student activity, we felt it desirable to apply each method to the same sample of student activity and then to assess the relative utility of the information provided by the two methods. We therefore elected to base our comparison on a series of video-recordings of student activity in laboratory classes.

Setting

The students involved in the study were first-year undergraduates taking four-year courses in agricultural and horticultural science at Lincoln College, New Zealand. In each of thirteen laboratory classes (five in zoology and eight in physics), we recorded the activity of a single student using video equipment and a small radio-microphone worn by the student. No attempt was made to choose students at random. Instead, the teachers supervising the classes were asked to provide us with a range of different types of student - able and not-so-able, industrious and not-so-industrious, gregarious and isolated, well-prepared and ill-prepared. A different student was used for each session. For a methodological study such as this, we felt it was more important

to record a wide range of different types of activity than to select a representative sample in the conventional sense.

In each case, we recorded the student's behavior for the entire period he or she was in the laboratory - between one and a half and three hours. Student participation in the study was voluntary.

We then used the video-recording in two distinct ways:

Method A: Interaction analysis

Within a day of recording each tape, we coded the video-recording of student behavior using an interaction analysis system specially developed for laboratory work by Kyle, Penick and Shymansky (1978). This system, known as SLIC (Science Laboratory Interaction Categories) - Student, involved coding student behavior at three-second intervals using ten categories of student behavior. A brief description of each of these categories is provided in Table 1. In addition, for each interval we also recorded whether the student was working alone or interacting with other students or with a demonstrator. If a student was interacting with other students, we recorded the number of students involved. The original SLIC system also required recording of the sex of the person with whom the student was interacting. We did not record this data as it seemed of little relevance to this study.

Table 1

Science Laboratory Interaction Categories (SLIC) - Student[a]

S - Shows

Manipulating materials with the purpose of transmitting information to teacher or student. Includes showing how to do something as well as demonstrating the result of doing something.

E - Manipulates apparatus, observes activity

Experimenting, manipulating equipment, observing experiment in progress. Other activity with lesson-related non-written materials.

T - Transmit information

Conveying ideas, thoughts, questions without manipulating equipment. Includes giving directions, feedback, answering questions or other behavior which transmit information. Includes rhetorical questions.

Q - Asks questions

Any attempt to gather information verbally from teacher or students.

L - Listens

Active behavior of receiving information from teacher or students.

O - Observes

Passive behavior allowing student to watch actions of teacher or students. Includes listening to conversations in which the subject is not directly involved.

R - Reads lesson-related material

Any reading related to the lesson or course goals. Includes notes that have been written by the student, notes written by another student, a textbook or a lab. manual.

W - Writes notes or records data

Note-taking, recording of data, answering of questions or other writing directly related to the lesson. Does not involve transmitting information immediately to another person.

M - Gets supplies, moves around the room

Moving around the room in a purposeful fashion, getting supplies, putting away materials or other mobile activities which facilitate the lesson.

Z - Non-lesson-related behavior

Behavior which does not facilitate the goals of the particular course or lesson, including interactive as well as non-interactive behaviors.

[a] Shymansky and Penick, 1979, p 198

Obviously, this produced a substantial set of data for each student. We chose to summarise this data in the form of tables and bar-graphs. Thus for each students we prepared:

1. A table of the relative frequency of behavior categories

This table allowed inspection of the relative proportion of time the student spent engaged in each of the ten categories of behavior. Summaries were provided for the entire period of the class and for each 20-minute interval. The table thus provided both an overall summary of student behavior and an indication of how behavior patterns changed during the period of the class.

2. A table of the relative frequency of recipient categories

This table presented the relative proportion of time the student spent alone, interacting with other students or interacting with a demonstrator. Again, summaries were provided for the entire period of the class and for each 20-minute interval to allow inspection of changes during the period of the class.

3. A table cross-tabulating behavior categories with recipient categories

This table allowed each behavior category to be related to each recipient category. Data for the total time period only was presented in this table.

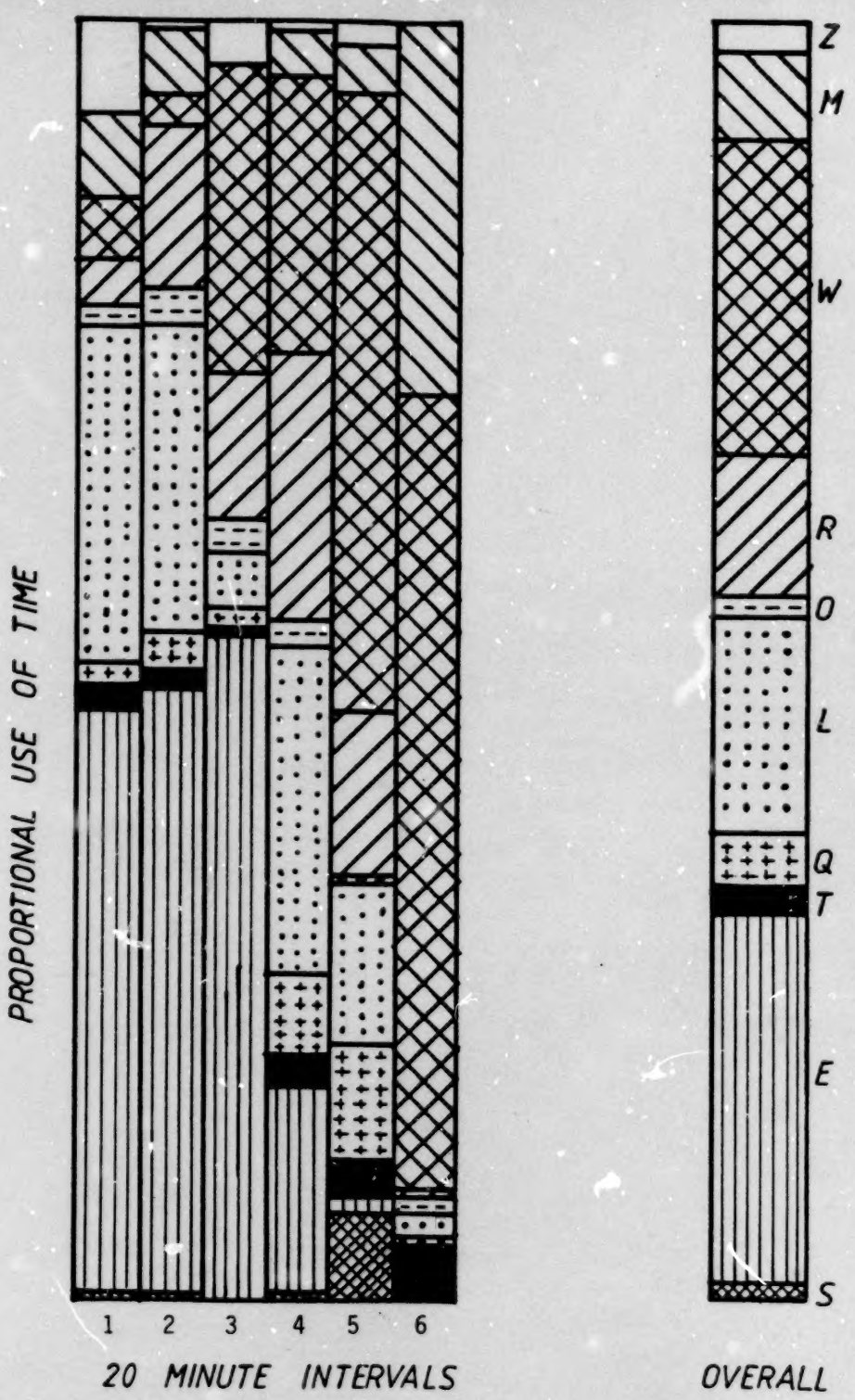
4. Bar-graphs of the relative frequency of behavior categories

These bar-graphs presented the same information as the first table. An example of a set of bar-graphs is presented in Figure 1.

Method B: Stimulated recall

Within three days of making each video-recording, we conducted an individual interview with the student involved. After a brief introduction explaining the purpose of the study and assuring confidentiality, we showed the student an excerpt from the beginning of the video-recording and asked the student to describe what he or she was doing at that time. With the video still running, we encouraged the student to expand on any issues raised in this description. We showed each student several excerpts in this way and used the same procedure to encourage the students to discuss their activity in the class. In other words, we used excerpts from the video-recording to stimulate the student's recall of activity in the class.

The interviews were unstructured in the sense that we used no schedule of set questions but instead allowed the student to discuss any issue triggered by viewing the video-recording. Because of this, some issues were raised in some interviews and not in others and no two interviews developed in the same way. The students often went on to discuss aspects of their courses other than the laboratory program and we sometimes initiated the replay of another excerpt from the video-recording to refocus their attention on their activity in the class. This helped to focus the interview and we also directed the interview to a limited extent by ensuring that every student discussed each of the following topics at some point in the interview:



LABORATORY WORK STUDY No.3.

FIGURE 1 : Bar-graphs of the relative frequency of behaviour categories

1. preparation for the lab. class
2. activity during the lab. class
3. writing the lab. report
4. aims of the lab. class
5. educational value of the lab. class

Each interview lasted between 30 and 60 minutes and, with the consent of each student, was recorded on audio-tape.

After completing all thirteen interviews, we analyzed the content of these audio-recordings and prepared a two to three page report on student responses from each interview. This report summarized student responses in the five areas listed above plus any other major issues raised by the student in the interview. An excerpt from one of the summaries is provided in Table 2.

Table 2 Excerpt from an interview summary

2. Activity during the lab.

He stated that the instructions in the lab. manual for the first exercise (wiring a circuit) did not appear to make sense. The student did not understand the procedure or method; it was not as he expected it (the circuit) to be and the instructions in the manual were not clear. On a second reading of the manual, he understood vaguely what was required, but he still felt slightly confused. The other students did not seem able to understand the instructions either. Although the exercises themselves were easy, this difficulty in interpretation of instructions made the exercise somewhat difficult. The student also found the concept of errors quite difficult, as he had not had to work through similar exercises for about two years.

Evaluation of the two methods

The main focus of this study was on the relative utility of the information produced by the two methods. As judges of the utility or applicibility of this information, we chose those most likely

to use or apply this information, the teachers of the courses involved. For the zoology course, the members of staff who had an on-going involvement with the laboratory program were two senior lecturers (professors) and three tutors (tenured teaching assistants). Similarly, two senior lecturers, one tutor and one senior technician were directly involved with the physics course. All nine members of staff agreed to participate in this study.

Each teacher was provided with a brief introduction to the project and two sets of summaries (Method A and Method B) for each of the students studied in their course (five for zoology and eight for physics). They were also provided with a summary of the SLIC - Student categories to aid their interpretation of the Method A summaries. In the introduction, the participants were advised that they were being asked to compare the quality of information provided by the two techniques in terms of its usefulness in aiding development of their laboratory program. In particular, we asked them to consider the following three questions:

- Do these methods provide you with useful information about student activity in the laboratory program?
- If you could only be provided with information from one method, would you prefer that from Method A (interaction analysis) or Method B (stimulated recall)? Why?
- What are the main limitations of the information provided by each method?

Within a week of issuing this material, we interviewed each of the teachers to obtain their answers to these questions and any other comments they wished to make. Each interview lasted between 20 and 60 minutes and teacher responses were recorded in written form by the interviewer.

RESULTS

All nine of the teachers interviewed declared the information provided them to be potentially useful in the development of their laboratory programs. As it happened, both sets of teachers were seriously considering a major review of their current laboratory programs and they expressed great interest in the information provided, even though we stressed that it was not intended to be representative of the whole class. None of those interviewed seriously questioned the validity or credibility of the information. Most commented that the general findings were consistent with their perception of student activity in the lab. classes.

The summaries therefore had a type of face validity or credibility which no doubt accounted for their apparent acceptance by the teachers. Several noted that they found it useful to have both Method A (interaction analysis) and Method B (stimulated recall) data side-by-side both because one confirmed the other and because the two methods produced complementary information. In particular, several noted that they often used information provided in the interview summaries to interpret aspects of the interaction analysis data.

Method A: Interaction analysis

Those interviewed varied considerably in their assessment of the usefulness of the information provided by Method A. Only two teachers rated this information as very useful for course development and most expressed distinct reservations about its utility. Some commented that they found the analysis 'interesting'

but could not imagine how they could use the information generated to improve their courses. Comments such as these were common:

'Method A provides a very useful back-up and quite a different sort of information... They're interesting results but how you could use them in making the labs better, I'm not quite sure.'

'I don't know what this behavior means. How can you use this information.'

All but one of those interviewed stated that Method A provided them with highly predictable information. They declared that it was mainly useful for verifying conclusions they had already made on the basis of informal observation in the laboratory:

'Method A gave me what I seemed to know already.'

'Method A confirmed what I'd already gathered about the labs.'

When pressed to specify what useful information Method A had provided them with, most found it difficult to give concrete examples. Three observed that they found it useful to learn how much time students spent on task and noted (with some satisfaction) that little time was coded as 'Z' (non-lesson-related behavior):

'It was useful to learn whether we were working them hard or not.'

'If there are great hunks of Z, then you've got a problem.'

Others singled out the proportion of time spent in reading (R) and writing (W) compared with that spent in manipulating apparatus and observing activity (E).

'The large amount of time spent reading and writing and less time spent manipulating, I find disturbing!'

'They spend a great deal of time "E-ing" and that's the avowed intent of the lab. - to get them to spend as much time as possible with hands on.'

Five teachers said that they found it useful to learn how much time the students spent working alone. All of the lab. classes studied required students to work alone and the analysis showed that all of the students studied/worked alone most of the time. However, several teachers noted that this could have been a measurement artefact as they had noticed that other students tended to avoid the student under observation. Most commented that the method used was a very reactive one which probably modified student behavior considerably.

In general, the teachers were quite critical of the information provided by Method A. Their main criticism was that it offered them a very limited type of information and much of this was already known to them. The category system used was often declared to be too crude to be really useful. For example, three teachers would have preferred to see the E category (manipulating apparatus, observing activity) broken down into sub-categories. Others queried whether it was really informative to know how much time a student spent staring at a lab. manual (R - reading) and declared themselves much more interested in knowing whether a student was, in fact, carefully reading instructions for the first time, re-reading them with growing exasperation or perhaps just staring through the book while

struggling mentally with some results just obtained. Some teachers also commented that the emphasis on observable behavior above ignored both student thinking and student feelings, two aspects of the process of study in laboratories which they felt were crucial to an understanding of student activity. As one teacher commented:

> *'They [Method A data] don't tell you much. They tell you what's going on in the lab. but you need to know much more than that... The human element is missing. You might just as well have been recording a flock of sheep. You can observe things but you don't know why they're going on.'*

Method B: Stimulated recall

The teachers interviewed were much more enthusiastic about the potential of Method B as a means of collecting useful data for curriculum review. All respondents endorsed the information generated by Method B as very useful for this purpose and claimed that even this limited methodological study provided them with information which they would use to modify the program in the next academic year.

The teachers noted that the information provided was not only directly useful but also interesting or even 'fascinating':

> *'The information I found most stimulating was the stimulated recall data.'*
>
> *'Method B was far more useful, but I'm probably a bit biased because it's more interesting.'*

When asked why they found Method B data valuable, they emphasized that its main value lay in providing them with a perspective of laboratory work from the viewpoint of the student:

> *'How the student was coping with the course. How he viewed the course and what he was getting out of it.'*
>
> *'You need to know about the quality of the interaction and this gives it to you.'*

The teachers said that Method B provided them not only with useful information about what students did in the laboratory but also about student attitudes to laboratory work. Several also noted the students not only talked about activity in the laboratory but also about lab-related activities outside of the formal class period.

When they were asked for specific examples of items of information which were particularly useful, the teachers most commonly cited information about student confusion about the aims and procedures of the laboratory exercises. In particular, student descriptions of their difficulties in interpreting the lab. manual were rated as especially useful. Three teachers stated that they found student comments about the relationship between the content of the laboratory program and the lecture program most useful. All of those interviewed valued feedback on how students felt about the lab. classes:

> *'How much they get out of the lab. depends on their attitudes. Why are they hostile?... They don't seem to get much out of it."*

'Because it gave information on student attitudes and their perception of what was going on.'

Some teachers expressed surprise that the students often expressed feelings of insecurity about their performance in laboratory classes.

When asked about the limitations of the information generated by Method B, most of those interviewed noted that the information was impressionistic and uneven. They observed that this was an inevitable outcome of the relatively unstructured interviewing technique used. Nonetheless, when asked whether the lack of structure in the interview presented problems in interpretation, they claimed that it did not. Nearly all of the teachers strongly favoured the unstructured approach and stated that they believed the imposition of greater structure would have resulted in a more systematic collection of data but that the information generated would have been less valid and less useful. Only one teacher argued for the advantages of more structured interviewing. Two noted that the use of the method with a larger sample of students might demand more directive interviewing and/or systematic content analysis of interview transcripts. Four lecturers stated that they would have preferred to have had some background information about each student so that they could interpret the summaries with that knowledge.

Method preference

Every teacher interviewed expressed a preference for the information generated by Method B (stimulated recall):

'Method B was far and away the most useful - the main reason is that it gives us the type of information I thought we needed.'

'I can see the advantages of both methods but I certainly prefer the second method.'

When asked to justify their preference, they reiterated the advantages of Method B reported above - that it provided them with detailed information about the quality of student activity in the laboratory and in lab.-related activity outside of the class. Four of the teachers observed that it provided them with the sort of information which they themselves would find difficult to obtain from the students. On the other hand, Method A could be performed by anyone:

'We couldn't get that sort of information out of the students outselves... We could do Method A ourselves - that's mechanical.'

Several expressed some surprise that they had found Method B data more useful than that from Method A. They noted that this was an unexpected finding which ran counter to their experience as researchers in the natural sciences:

'It's probably silly but one has been taught to place more faith in this [Method A data] than this [Method B data] but I actually find Method B more easy to relate to the lab.'

BEST COPY AVAILABLE

DISCUSSION

This group of teachers found the information generated by the stimulated recall interviews far more useful than that generated by the SLIC - Student interaction analysis. In very general terms, they found the interaction analysis data coarse-grained, of low quality and predictable. In short, for their purposes it had relatively low explanatory power in helping them to understand how students learn in the laboratory. On the other hand, the stimulated recall interview data generated information which they found more detailed and of higher quality. They claimed that it provided them with new insight into student activity in laboratories. They favoured the second method even though it was, as they saw it, less rigorous than interaction analysis in that it produced information that was largely anectodal and idiosyncratic.

Part of the reason for this clear distinction probably lay in the relative accessibility of the information. Method A data was quantitive and presented in tables and figures with no verbal interpretation. It was therefore a somewhat daunting mass of figures and it was declared somewhat indigestible by some respondents. Nonetheless, each of those interviewed was familiar with the interpretation of this type of data as they were all active researchers in their own highly quantitative natural science fields. In fact, they all stated that they had little difficulty in interpreting the tables and figures. Their main complaint was that, having done so, they found the effort hardly worthwhile in terms of the usefulness of their conclusions. On the other hand, none of them had any experience of research in the social sciences and of the interpretation of interview data.

They sometimes expressed surprise at the, to them, counter-intuitive finding that the somewhat messy qualitative description was more convincing and provided them with more useful information than the neater quantitative approach based on direct observation, the method commonly used in their own research.

One of the major challenges facing the evaluator is to provide information which the client finds both credible and useful. Cronbach et al. (1980) emphasise that for an evaluation to be effective, it must provide information which has 'leverage' ie. has the potential to change the course of events:

> 'The evaluator is indulging himself if he deliberately invests in information that has little leverage. Social scientists are tempted to study the questions to which their research technology is suited and the questions that have theoretical implications rather than to study what matters most to the client. An evaluation is unlikely to carry much weight if it fails to address questions that have great leverage.'
> (Cronbach et al., 1980, p 265)

It is clear that the evaluator must use methods which are both credible and produce information useful to the client.

If the instructors interviewed in this study commissioned a formative evaluation of their laboratory program, the results of a series of stimulated recall interviews would be likely to have much more leverage than the results of a SLIC - Student interaction analysis. However, they may not have been able to

tell us this before being presented with examples of the type of information produced by the two methods. In our experience, most college instructors find it difficult to spell out what information they want collected in an evaluation study and they also have little idea of what type of information is generated by different methods of inquiry. In this study, we asked our potential clients to evaluate the effectiveness of two methods of collecting information. This is a step which is rarely included in small-scale curriculum evaluation projects although it may be included in some larger projects (Cronbach et al, 1980). It has certainly been advocated as an important part of naturalistic evaluation (Guba and Lincoln, 1981). As a result of this study, we intend to include such a step in future projects. We have learnt that methods which are quick and dirty in terms of traditional research criteria may produce data with more leverage than those which are slow and methodologically pure.

But this study also raises some more fundamental questions. In many ways, the two techniques we chose to compare are examples of two very different research paradigms. One, interaction analysis, is an example of a scientific or positivist approach to social inquiry. The other, stimulated recall, was aimed very much at exploring how the student interpreted life in the laboratory. It was designed to encourage the students to tell us what meanings the laboratory class had for them. As such, it was an example of a more naturalistic or interpretive approach. Guba and Lincoln (1981) have suggested that rigor is associated with the scientific paradigm and relevance with the naturalistic paradigm. Several commentators, including Schwandt (1982), have pointed out that

this is a false dichotomy. However, the teachers we interviewed did describe the interaction analysis as rigorous but lacking relevance and the stimulated recall interviews as lacking rigor but producing highly relevant information. We found it fascinating to hear these instructors, themselves positivist researchers in the natural sciences, make comments about the two methods which seemed to echo the arguments made by critics of a positivist approach to social inquiry. The instructors, in their various ways, suggested that their criticism of the data produced by SLIC - Student could also be extended to other strictly observational systems of inquiry. They stated that they thought it unlikely that any analysis of overt behavior could provide them with the sort of information they needed to understand student learning in laboratories. Many of them went on to state that a knowledge of how students make meaning of laboratory work seemed more likely to lead to an understanding of how to improve learning in the laboratory.

The arguments the instructors used had a familiar ring to them. They were similar to those used by critics of the positivist paradigm such as Parlett and Hamilton (1972) and Guba and Lincoln (1981). As one of our interviewees said:

> *'I'm not too interested in their actual behavior in labs. Each of them will interpret it in different ways and that's what I want to know.'*

Perhaps he is right. Perhaps any study of behavior alone is not only inappropriate for quick and dirty evaluation studies but also so limited that it is unlikely to lead to fundamental understanding of learning in the laboratory.

REFERENCES

Brinkerhoff, R.O., Brethower, D.M., Hluchyj, R., and Nowakowski, J.R. (1983). Program evaluation: a practitioner's guide for trainers and educators. Boston: Kluwer-Nijhoff.

Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., Walker, D.C., and Weiner, S.S. (1980). Toward reform of program evaluation. San Francisco: Jossey-Bass.

Guba, E.G. and Lincoln, Y.S. (1981). Effective evaluation. San Francisco: Jossey-Bass.

Kyle, W.C., Penick, J.E. and Shymansky, J.A. (1978). Science Laboratory Interaction Categories (SLIC). Technical Report 15. Iowa City: Science Education Centre, The University of Iowa.

Ogborn, J. (Ed.) (1977). Practical work in undergraduate science. London: Heinemann.

Parlett, M., and Hamilton, D. (1972) Evaluation as illumination: a new approach to the study of innovatory programmes. Occasional Paper No. 9. Edinburgh: Centre for Research in the Educational Sciences, University of Edinburgh.

Power, C. (1977) A critical review of science classroom interaction studies. Studies in Science Education, 4, 1-30.

Schwandt, T.A. (1982). Rigor and relevance in vocational education evaluation: an explication and review. Journal of Vocational Education Research, 1, 1-14.

Shymansky, J.A., and Penick, J.E. (1979). Use of systematic observations to improve college science laboratory instruction. Science Education, 63, 195-203.